(54) Title: PROTECTING NUCLEIC ACIDS AND METHODS OF ANALYSIS

(57) Abstract

A selected region of a nucleic acid is protected from attack by nuclease by complexing to the nucleic acid a nucleic acid analogue of the PNA type. The surviving sequence may be detected in an assay, optionally after amplification. A PCR reaction may be sterilised and its product assayed by protecting a characterising region of the product by PNA hybridisation followed by nuclease degradation of unprotected nucleic acid.

-1-

PROTECTING NUCLEIC ACIDS AND METHODS OF ANALYSIS

The present invention relates to methods of protecting
nucleic acids from attack, eg. by nucleases, and analytical
5   methods and in particular to methods for detecting the
presence in a nucleic acid containing sample of a sequence of
interest.

At present the usual way of determining whether a
particular base sequence is present in a sample of nucleic
10   acid is filter hybridisation. The nucleic acid is extracted
from starting materials and purified. It may need to be
amplified by a process such as PCR. The nucleic acid is then
immobilised in single stranded form on a filter and is probed
with a labelled oligonucleotide having a sequence comple-
15   mentary to that which is to be detected. Before being
immobilised, the nucleic acid may be run on a gel to separate
different molecular weight nucleic acid fragments. This is
the widely used Southern blot procedure. This procedure has
numerous drawbacks. The nucleic acid needs to be separated
20   from other materials. Where the nucleic acid is DNA, it will
normally be necessary to go through a denaturation step. The
stringency conditions of the hybridisation will need careful
adjustment to avoid false positives and to achieve the desired
level of discrimination between similar sequences. Because
25   of the limited signal to noise of the filter hybridisation,
amplification steps will often be necessary to produce
sufficient of the sequence of interest for detection. This
generates many opportunities for the production of false
negatives.

30   Nucleic acid analogues having important new utilities in
assay procedures and in the field of diagnostics have been
described in WO 92/20703. These nucleic acid analogues had
a number of new properties making them of special importance
in the field of diagnostics as well as in the field of anti-
35   sense therapeutics. Such nucleic acid analogues are referred
to herein as "PNAs".

-2-

They typically feature a polyamide backbone bearing a
sequence of ligands which are nucleic acid bases. The
analogues described there have the property of hybridising
with great specificity and stability to natural nucleic acids
5   of complementary sequence.

We have now discovered that such nucleic acid analogues
have the further valuable property of protecting sequences of
single stranded nucleic acids to which they are hybridised
from being degraded by certain reagents under circumstances
10 and conditions in which the reagents would normally be
effective to produce degradation of said single stranded
nucleic acids. Regions of the nucleic acid not hybridised to
the nucleic acid analogue may be degraded as normal.

Accordingly, the present invention provides a method of
15 protecting a selected region of a nucleic acid from attack by
a reagent, which method comprises forming a complex between
the nucleic acid and a nucleic acid analogue which hybridises
thereto in a sequence selective manner and exposing said
nucleic acid to a nucleic acid attacking reagent, wherein said
20 complex is more stable against attack by said reagent than
said starting nucleic acid.

The nucleic acid may be RNA or DNA.

Preferably, said reagent is a nuclease and is capable of
degrading said nucleic acid either completely or by cleaving
25 it at certain locations.

In the most straightforward case, the selected region of
the nucleic acid to be protected will be constituted by the
nucleic acid sequence to which the nucleic acid analogue
hybridises. However, where the attacking reagent is an
30 exonuclease capable of degrading nucleic acids under the
conditions used only by working in from the 5' and 3' ends,
a larger selected region may be protected by placing a
respective nucleic acid analogue at each end of the selected
region leaving part of the selected region unhybridised
35 between the two nucleic acid analogue sequences. The
intervening part of the selected region is fenced off from

-3-

access by the exonuclease. Where the exonuclease is, under
the conditions employed, able only to attack from the 5' or
3' end, a single nucleic acid analogue sequence may be
employed to fence off all that part of the nucleic acid which
5   lies 3' or 5' respectively to the nucleic acid analogue.

Where an endonuclease is employed as the attacking
reagent, a short stretch of nucleic acid exposed unhybridised
between two nucleic acid analogue oligomers lying adjacent one
another on the nucleic acid may none the less form part of the
10 selected region which is protected, by virtue of some
interaction between the end bases of the two nucleic acid
analogue sequences or because of interference by the nucleic
acid analogue with the activity of the endonuclease.

The nucleic acid analogue is preferably one comprising
15 a polymeric strand which includes a sequence of ligands bound
to a backbone made up of linked backbone moieties, which
analogue is capable of hybridisation to a nucleic acid of
complementary sequence.

Said nucleic acid analogue backbone is preferably a
20 polyamide, polythioamide, polysulphinamide or polysulphonamide
backbone.

Preferably, said linked backbone moieties are peptide
bonded amino acid moieties.

The nucleic acid analogue is preferably capable of
25 hybridising to a nucleic acid of complementary sequence to
form a hybrid which is more stable against denaturation by
heat than a hybrid between the conventional deoxyribo-
nucleotide corresponding in sequence to said analogue and said
nucleic acid.

30 Preferably, said nucleic acid analogue is a peptide
nucleic acid in which said backbone is a polyamide backbone,
each said ligand being bonded directly or indirectly to an aza
nitrogen atom in said backbone, and said ligand bearing
nitrogen atoms mainly being separated from one another in said
35 backbone by from 4 to 8 intervening atoms.

-4-

Preferably also, nucleic acid analogue is capable of hybridising to a double stranded nucleic acid in which one strand has a sequence complementary to said analogue, in such a way as to displace the other strand from said one strand.

5   Preferably, the nucleic acid analogue has the general formula 1:
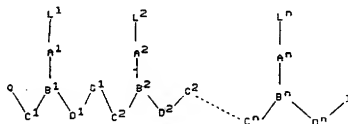
Formula 1

wherein:

n is at least 2,

10   each of L^1-L^n is independently selected from the group consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally occurring nucleobases, non-naturally occurring nucleobases, aromatic moieties, DNA intercalators, nucleobase-binding groups, heterocyclic moieties, and reporter ligands, but

15   normally at least one L will be a nucleobase binding group such as a naturally occurring nucleobase and preferably at least 90% of the groups L will be such nucleobase binding groups;

each of C^1-C^n is $(CR^6R^7)_y$ where R^6 is hydrogen and R^7 is

20   selected from the group consisting of the side chains of naturally occurring alpha amino acids, or R^6 and R^7 are independently selected from the group consisting of hydrogen, $(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-$C_6)$alkoxy, $(C_1$-$C_6)$alkylthio, NR^3R^4 and SR^5, where R^3 and R^4

25   are as defined below, and R^5 is hydrogen, $(C_1$-$C_6)$alkyl, hydroxy, alkoxy, or alkylthio-substituted $(C_1$ to $C_6)$alkyl or

-5-

R^6 and R^7 taken together complete an alicyclic or heterocyclic system;

each of D^1-D^n is $(CR^6R^7)_z$ where R^6 and R^7 are as defined above;

each of y and z is zero or an integer from 1 to 10, the sum y + z being from 2 to 10 (preferably being more than 2, and most preferably each of y and z being 1 or 2;

each of G^1-G^{n-1} is $-NR^3CO-$, $-NR^3C^S-$, $-NR^3SO-$ or $-NR^3SO_2-$, in either orientation, where R^3 is as defined below:

each of A^1-A^n and B^1-B^n are selected such that:

(a)   A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or R^3N^+; or

(b)   A is a group of formula (IId) and B is CH;

Formula IIa

Formula IIb

Formula IIc

Formula IId

-6-

wherein:

X is O, S, Se, NR^3, CH_2 or $C(CH_3)_2$;

Y is a single bond, O, S or NR^4;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, (C_1-C_4)alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

each R^3 and R^4 is independently selected from the group consisting of hydrogen, (C_1-C_4)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C_1-C_4)alkyl, hydroxy, alkoxy, alkylthio and amino;

Q is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NR'R''$, where R' and R'' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphosphates, oligonucleotides, including both oligoribonucleotides and oligodeoxyribonucleotides, oligonucleosides and soluble and non-soluble polymers.

More preferably, said nucleic acid analogue comprises a compound of the general formula III, IV or V:

-7-

Formula III

Formula IV

5

10

-9-

Formula V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally-occurring nucleobases, and non-naturally occurring nucleobases;

each R^7 is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

R^h is OH, NH_2 or $-NHLysNH_2$; and

R^i is H or $-COCH_3$.

The method of protecting a nucleic acid described above can be used in a method of detecting a nucleic acid sequence. Thus the present invention includes in a second aspect a method of detecting the presence in a nucleic acid sample of a sequence, which method comprises exposing the sample of nucleic acid to a nucleic acid analogue capable of hybridising in sequence selective manner to said sequence if present in

-9-

said nucleic acid sample under hybridising conditions so as
to form a complex between said nucleic acid analogue and a
region of said nucleic acid containing said sequence, exposing
said nucleic acid to a reagent capable of degrading said
5   nucleic acid under conditions such that the said region of
said nucleic acid forming said complex is protected from
attack by said reagent whilst the remainder of said nucleic
acid is degraded, and detecting the presence of said complex.

There will be many ways of detecting the presence of said
10   complex. These may be divided into those which detect the
complex as such and those in which the complex is broken up
into its constituents of nucleic acid and nucleic acid
analogue and the presence of the nucleic acid or the nucelic
acid analogue is detected. The existence of the complex can
15   be deduced from the survival of the nucleic acid.

In the first group of methods, reliance may be placed
upon differences in properties of the nucleic acid analogue
and the nucleic acid/nucleic acid analogue hybrid. Thus
preferably, the presence of said complex is detected by
20   comparing the movement of said complex in electrophoresis with
the movement of said nucleic acid analogue under similar
conditions.

Preferably, the nucleic acid analogue or the nucleic acid
carries a detectable label, suitable labels for the nucleic
25   acid analogue include a fluorescent label, a radio-label, an
enzyme label, biotin, a spin label, a chemiluminescent label,
an antigen label or an antibody label. The most preferred of
these are the use of a fluorophore or a radioisotope as a
label but generally any method of labelling applicable to
30   peptides can be used.

In a preferred practice of the second aspect of the
invention, the nucleic acid analogue is a labelled nucleic
acid analogue produced by a method for labelling a nucleic
acid analogue comprising providing a nucleic acid analogue
35   with a peptide motif capable of functioning as a substrate for
an enzyme in a labelling reaction and carrying out a said

labelling reaction comprising reacting the peptide motif of the nucleic acid analogue under the influence of an enzyme with a source of said label.

Preferably, said label is a radio-label.

The source of said label is preferably radio-labelled ATP.

Said enzyme is preferably a protein kinase.

The peptide motif is preferably the kemptide motif.

Preferably, the labelling reaction is a phosphorylation at a serine residue of said peptide motif.

An alternative method of labelling a nucleic acid analogue involves providing, preferably at one end of the nucleic acid analogue, a chelating moiety capable of binding at least one metal ion by chelation and chelating a radioisotope or fluorophore directly or indirectly thereto. For instance, a europium ion may be chelated by a suitable chelating moiety carried by the nucleic acid analogue to act as a fluorescent label.

An alternative detection method in this first group makes use of the nucleic acid sequence which is protected to act as a means of linking two nucleic acid analogue sequences which would otherwise not associate. If each bears a different type of label, the association of the two labels can be detected. By way of example, two nucleic acid analogue oligomers may be hybridised to respective ends of a target nucleic acid sequence forming part of a larger nucleic acid sequence. One nucleic acid analogue is labelled with a moiety suitable for linking to a solid phase e.g. a chelating peptide motif such as His$_5$. The other is provided with a detectable label, e.g. is radio-labelled. After digestion of unprotected nucleic acid, the two nucleic acid analogue sequences are left joined via hybridisation to the target nucleic acid protected sequence. The complex may be captured on a solid support bearing chelatable nickel ions via chelation by the first nucleic acid analogue. Radio-labelled excess second nucleic acid analogue may be washed away and the radio activity of the

-11-

second labelled nucleic acid analogue may be detected on the
solid support or after elution from it.

Other examples would include having one nucleic acid
analogue bearing biotin as a moiety suitable for linking to
a solid phase instead of the chelating peptide motif and the
second nucleic acid analogue linked to a detectable label,
e.g. an enzyme or a fluorophore. If the detectable label is
an enzyme capable of catalysing the production of a detectable
product, the assay is analogous to the well established ELISA
principle in immunoassay.

Alternative means for separating hybridised from non-
hybridised nucleic acid analogue include reverse phase or ion
exchange chromatography, or gel filtration. An antibody
selection procedure could be used employing an antibody with
specificity for intact nucleic acid or with specificity for
the nucleic acid/nucleic acid analogue hybrid.

In the second group of detection methods, the complex
between the nucleic acid and the nucleic acid analogue is
denatured and the survival of the protected nucleic acid
sequence is demonstrated, suitably by known nucleic acid
sequence detection methods such as a hybridisation assay.
Optionally, the nucleic acid sequence may be subjected to an
amplification procedure. LCR will generally be preferred for
amplifying short protected sequences and PCR for longer ones.

Nucleic acid amplification procedures, especially the
well known PCR procedures often encounter difficulties when
the target sequence to be amplified is present amongst a large
amount of other nucleic acid.

As is shown above, the present invention provides a ready
means of "cleaning up" a nucleic acid sample by degrading all
nucleic acid present except the target sequence, thus allowing
amplification to be more easily carried out whether for the
purposes of assay or preparatively.

Accordingly, the present invention provides in a third
aspect, a method of carrying out a nucleic acid amplification
comprising protecting a selected region of a nucleic acid

within a nucleic acid sample by hybridising to said nucleic acid a nucleic acid analogue which hybridises thereto in a sequence specific manner, exposing said sample to a nucleic acid attacking reagent to degrade the nucleic acid in the
5   sample except for said selected region, and amplifying said selected region.

To free the protected sequence for amplification, the nucleic acid/nucleic acid analogue hybrid may be denatured by heat. Any amplification method may be employed, including all
10 presently known method such as PCR and its variants, LCR and 3SR.

In WO   (PCT/EP93/01435) a method is described for sterilising a nucleic acid amplification by hybridising to the amplification product a nucleic acid analogue which
15 will prevent any amplification from acting as an amplifiable target. This may be still further improved according to a fourth aspect of the present invention by protecting a region of the amplification product which excludes at least one primer binding site necessary for repetition of the
20 amplification by a method in accordance with the first aspect of the invention, degrading the remaining nucleic acid, including the primer binding sites of the amplification product, and detecting the survival of the protected nucleic acid as discussed above. Preferably, at least some, more
25 preferably the whole, of all the primer binding sequences are degraded.

The nucleic acid analogue may be present throughout a PCR procedure, provided that the temperatures employed are high enough to melt the nucleic acid analogue off the nucleic acid,
30 or may· be added at the end of the amplification. The attacking reagent will need to be added following the amplification but may be added from a separate compartment within a closed amplification apparatus so that amplified product never has an opportunity to escape and become a
35 contaminant.

-13-

        The invention will be illustrated by the following
examples with reference to the accompanying drawings in which:
        Figure 1 is an autoradiograph of a gel showing the
results of Examples 3 to 5.

5   One convenient method of producing a PNA bearing a radio-
label of high specific activity particularly preferred for use
in the present invention is described in our co-pending
British patent application bearing agent's reference P3685GB,
filed on the same day as this application. According to such

10   a method in its preferred form, a PNA is prepared having a
terminal kemptide extension which undergoes a reaction with
^{32}P ATP mediated by protein kinase A to acquire a ^{32}P label
bound to a serine residue.   PNA's having peptide extensions
of desired amino acid sequences can conveniently be produced

15   by the Boc or Fmoc solid phase techniques well understood in
the art once a starting PNA sequence has been built-up on the
solid support suitably by using the Boc solid phase synthesis
described in WO92/20703.

        According to preferred forms of the first aspect of the
20   present invention, a nucleic acid containing sequences to be
protected may be treated with a PNA having a sequence
complementary to the sequence of interest. If the nucleic
acid is DNA, it may be in double stranded form and one may
rely on the ability of the nucleic acid analogue to hybridise

25   to its target sequence in double stranded form. However, it
may be necessary and will normally be desirable to denature
the double stranded target DNA, usually by heating. The DNA
may be quite impure, i.e. other cell products may be present.
Also, the DNA may be whole genomic DNA or total cell RNA

30   extracted from cells and not digested into shorter lengths.
The DNA or RNA could also be the product of a nucleic acid
amplification procedure such as LCR or PCR or a NASBA (3SR)
amplification. Preferred PNA's will bind sequence selectively
with high affinity, even if the target DNA is double stranded.

35   The nucleic acid sample may then be treated with a nuclease
or nuclease mixture to degrade all of the nucleic acid not

-14-

protected by hybridisation to the PNA. The PNA may be present
in substantial excess to ensure that if there is even one
correct sequence present in the nucleic acid sample, it will
be hybridised and protected.

5   The reaction product may be subjected to electro-phoresis
on a gel or in a capillary to separate PNA which is hybridised
to nucleic acid from non-hybridised PNA. Because PNA is an
essentially neutral molecule, having one positive charge per
molecule typically at a terminal amine group, it shows little
10   tendency to migrate during electrophoresis. Nucleic acids on
the other hand are highly negatively charged and hybridisation
of PNA to a nucleic acid fragment results in a negatively
charged hybrid which will migrate under electrophoresis in the
opposite direction to the PNA itself. This effects a ready
15   and highly specific separation of the hybridised and non-
hybridised PNA. If the PNA is appropriately labelled, it is
possible to detect extremely small quantities of it and the
analytical method according to the second aspect of the
invention provides an excellent signal to noise ratio.

20   PNA's are able to distinguish nucleic acid sequences
differing by only one residue. Accordingly, the methods
described above offer the possibility of a very quick and
simple assay capable of determining whether a sample of
genomic DNA contains a particular allelic variation within a
25   particular gene. Furthermore, as PNA's of different lengths
hybridised to nucleic acids are readily separable under
electrophoresis, the possibility also exists of probing a
nucleic acid sample with a plurality of PNA's simultaneously,
digesting away all the unhybridised nucleic acid and by
30   electrophoresis separating all the PNA/nucleic acid hybrids
from the unhybridised excess PNA'S and from each other. One
may thus in a single assay demonstrate that a DNA sample
contains, possibly on different genes, some particular
combination of allelic variations of interest, eg. one found
35   to give rise to a particular phenotype such as a disease
state.

-15-

Example 1

Preparation of a PNA-kemptide chimera.

5   The solid phase PNA synthesis described in
WO 92/20703 was used to build up the sequence:
Boc-NH(CH$_2$)$_5$CONH-TG(Z)T.A(Z)C(Z)G(Z).TC(Z)A(Z).C(Z)A(Z)
A(Z).C(Z)TA(Z)-CONH-resin

10   The N terminal Boc group was removed by treatment with
TFA and used as a starting point for a standard boc type solid
phase peptide synthesis of the kemptide motif via the linker
6-amino-hexanoic acid to produce the chimera:-

15   Boc-Leu-Arg(Tos)-Arg(Tos)-Ala-Ser-(Bzl)-Leu-Gly-NH(CH$_2$)$_5$CONH-
TG(Z)T.A(Z)C(Z)G(Z).TC(Z)A(Z).C(Z)A(Z)   A(Z).C(Z)TA(Z)-CONH-
resin

         The protection groups were removed and the product
20   cleaved from the resin by the Low-High TFMSA procedure.   The
raw product was purified by preparative HPLC (reversed phase
C$_{18}$ eluting with a gradient of A:0.1% TFA in water (MilliQ)
and B:0.1%TFA, 10% water,90% acetonitrile).   The purified
chimeric PNA-kemptide was characterized by analytical HPLC and
25   FAB-MS.

Example 2

Radio labelling of the kemptide motif (Leu-Arg-Arg-Ala-Ser-
30   Leu-Gly) by protein kinase A

         The PNA-kemptide chimera of the formula:-

         H-Leu-Arg-Arg-Ala-Ser-Leu-Gly-TGTACGTCACAACTA-NH$_2$
35   was labelled with ^{32}P in a reaction mixture containing:

-16-

```
PNA-kemptide, 10µM....................        5µl
10 x Protein Kinase A buffer..........        5µl
γ 32P ATP (>5000 Ci/mmol;50µCi/µl)....        10µl
Protein kinase A (Boehringer; 5mU/µl).        0.2µl
```
5   H_2O.................................   30µl

        The reaction was incubated for 30 minutes at 30°C and
    then for 10 minutes at 65°C before being centrifuged for 30
    seconds at 15000g. The supernatant was transferred to a new
10 eppendorff tube. Water was added to 1ml and the labelled PNA-
    kemptide was separated from unincorporated γ 32P ATP by anion
    exchange chromatography using a DEAE Sephadex A-50 anion
    exchange column.
        The specific activity of the PNA-kemptide was estimated
15 at 1×10^8 cpm/µg PNA-kemptide.

    Example 3. Detection of DNA sequence via protection by PNA.

        The 15-mer PNA bearing the radio-labelled kemptide motif
20 as made in Example 2 was incubated for 15 minutes separately
    with each of the following nucleic acids:

    1.   None (See Fig. 1 lane 1)
    2.   Complementary DNA 15-mer. (See Fig. 1 lanes 2 and 5)
25 3.   DNA 10-mer complementary to part of the sequence of the
    PNA. (See Fig. 1 lane 3)
    4.   DNA 40-mer 5'CTAGAGGATCTAGTTGTGACGTACAGGATCTTTTTCATAG-3
    containing a 15 base sequence complementary to the PNA in the
    centre of the sequence. (See Fig. 1 lanes 4 and 6)

30

        In each case, 4µl of a 10µM solution of the oligonucleo-
    tide was mixed with 5µl of the labelled PNA in a 10µl reaction
    volume containing 30mM NaAc (pH 5.0), 50mM NaCl, 10mM $ZnCl_2$
    and 55 %v/v glycerol. Incubation for 15 minutes took place
35 at room temperature.

-17-

In the case of one out of two reactions using the 40-mer DNA and in the case of one out of two reactions using the 15-mer DNA, 15 units of mung bean nuclease were added and the reactions were incubated at 37° for 5 minutes.

5   All reactions were terminated by adding 10μl of loading buffer (30% v/v glycerol,0.25% w/v bromophenol blue, 0.25% xylene cyanol, 0.01% SDS) and the reaction mixtures were subjected to electrophoresis in a 10% polyacrylamide gel and subsequently to autoradiography.

10   The results are shown in Figure 1 lanes 1 to 6. In lane 1 the labelled PNA is run alone and nothing is seen because it is immobile on the gel. In lane 2 the labelled PNA is hybridised to the 15-mer DNA and a band is seen for the hybrid. In lane 3 the labelled PNA is hybridised to the 10-

15   mer DNA and it can be seen that the band produced is clearly distinguishable from that in lane 2 in the distance run on account of the difference in the size of the two oligo-DNA's involved. In lane 4 the PNA is hybridized to the 40-mer DNA and once again the distance run by the band on the gel allows

20   one to discriminate between the ^{15}PNA/^{10}DNA and ^{15}PNA/^{15}DNA hybrids and the ^{15}PNA/^{40}DNA hybrid.

In lane 5 one sees the result of nuclease digestion in the case of the ^{15}PNA/^{15}DNA hybrid. The band is the same as that in lane 2 showing that the nuclease has been unable to

25   degrade the hybridised DNA. Lane 6 shows the result of the nuclease digestion of the ^{15}PNA/^{40}DNA hybrid. Once again, the band is as in lane 2, showing that only that part of the DNA which is hybridised to the PNA is protected against the nuclease. The internal 15-mer sequence in the 40-mer DNA

30   which is complementary to the PNA has therefore been detected by the procedure described above with reference to lane 6.

Example 4. Protection of RNA against nuclease by PNA.

35   PNA T$_{10}$ bearing the kemptide motif was labelled as in
Example 2. The labelled PNA was incubated with a mixture of

-18-

polyadenylated RNA's according to the following procedure.
PNA (5μl) was incubated with 4μl of 100ng/μl mixed in vitro
RNA transcripts (Gibco BRL) in a 10μl reaction volume
containing other constituents and under time and temperature
5   conditions as in Example 3. A similar incubated mixture was
further incubated with mung bean nuclease as in Example 3 and
the reaction mixtures were prepared for and subjected to
electrophoresis and autoradiographed as in Example 3. The
results are seen in lanes 7 and 8 of Figure 1.

10   Lane 7 shows an unresolved ladder of different sizes of
PNA/RNA hybrids. In lane 8, the nuclease has resolved the
ladder into a single band by trimming off all the RNA's to the
A_{10} portion protected by the PNA.

15   Example 5. Detection of single base pair variation in 15-mer
DNA

The 15-mer labelled PNA used in Example 3 was hybridized
to two different DNA's under both of two different temperature
20   conditions as follows.

The labelled PNA was hybridised to complementary 15-mer
DNA at incubation temperatures of 55° and 65°C respectively
and in each case the reaction mixture was treated at that
temperature with mung bean nuclease. In other respects the
25   conditions were as in Example 3. Two similar incubations took
place using the single base mismatched 15-mer DNA 5'-
TAGTTGCGACGTACA -3' at the same two temperatures with
subsequent addition of mung bean nuclease. The resulting
autoradiographs are seen in lanes 9 to 12 of Figure 1.

30   In lane 9 the result of the hybridization of the PNA to
its fully complementary DNA at 65°C with nuclease degradation
is a band corresponding to protection of the DNA by the PNA.
Lane 10 shows the result of hybridisation at 65°C with the
single base mismatch. The DNA is not protected and no hybrid
35   is seen. Lane 11 shows the result of the experiment at 55°C
with the fully complementary sequence and lane 12 shows the

-19-

result of the 55°C experiment with the single mismatch. In
both these lanes the protected DNA is detected.

   The melting point for the perfectly matched 15-mer hybrid
is 69°C. Under the stringent hybridisation conditions imposed
5 by operating close to this melting point, the PNA
distinguishes between the perfectly matched DNA and the single
base mismatch for which the melting point is 61°C. At 55°C
however, hybrids are formed both with the perfect match and
the single base mismatch as would be predicted from the
10 melting points of the hybrids.

   Whilst the invention has been described with reference
to the preferred practice thereof as illustrated by the above
examples, many variations and modifications are possible
within the scope of the invention.

-20-

CLAIMS

1.   A method of protecting a selected region of a nucleic
acid from attack by a reagent, which method comprises forming
a complex between the nucleic acid and a nucleic acid analogue
which hybridises thereto in a sequence selective manner and
exposing said nucleic acid to a nucleic acid attacking
reagent, wherein said complex is more stable against attack
by said reagent than said starting nucleic acid.

2.   A method as claimed in claim 1, wherein said reagent is
a nuclease.

3.   A method as claimed in Claim 1 or Claim 2, wherein the
nucleic acid analogue comprises a polymeric strand which
includes a sequence of ligands bound to a backbone made up of
linked backbone moieties, which analogue is capable of
hybridisation to a nucleic acid of complementary sequence.

4.   A method as claimed in Claim 3, wherein said nucleic acid
analogue   backbone   is   a   polyamide,   polythioamide,
polysulphinamide or polysulphonamide backbone.

5.   A method as claimed in Claim 4, wherein said linked
backbone moieties are peptide bonded amino acid moieties.

6.   A method as claimed in any preceding claim, wherein the
nucleic acid analogue is capable of hybridising to a nucleic
acid of complementary sequence to form a hybrid which is more
stable against denaturation by heat than a hybrid between the
conventional deoxyribonucleotide corresponding in sequence to
said analogue and said nucleic acid.

7.   A method as claimed in any preceding claim, wherein said
nucleic acid analogue is a peptide nucleic acid in which said
backbone   is   a   polyamide   backbone,   each   said   ligand   being

-21-

bonded directly or indirectly to an aza nitrogen atom in said
backbone, and said ligand bearing nitrogen atoms mainly being
separated from one another in said backbone by from 4 to 8
intervening atoms.

5

8.   A method as claimed in any preceding claim, wherein the
nucleic acid analogue is capable of hybridising to a double
stranded nucleic acid in which one strand has a sequence
complementary to said analogue, in such a way as to displace
10   the other strand from said one strand.

9.   A method as claimed in any preceding claim, wherein the
nucleic acid analogue has the general formula 1:
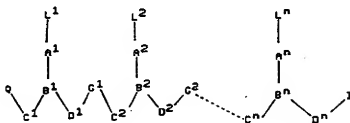
Formula 1

wherein:

15   n is at least 2,
        each of L^1-L^n is independently selected from the group
consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally
occurring nucleobases, non-naturally occurring nucleobases,
aromatic moieties, DNA intercalators, nucleobase-binding
20   groups, heterocyclic moieties, and reporter ligands;
        each of C^1-C^n is $(CR^6R^7)_y$ where R^6 is hydrogen and R^7 is
selected from the group consisting of the side chains of
naturally occurring alpha amino acids, or R^6 and R^7 are
independently selected from the group consisting of hydrogen,
25    $(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-
$C_6)$alkoxy, $(C_1$-$C_6)$alkylthio, NR^3R^4 and SR^5, where R^3 and R^4

-22-

are as defined below, and R^5 is hydrogen, (C_1-C_6)alkyl,
hydroxy, alkoxy, or alkylthio-substituted $(C_1$ to $C_6)$alkyl or
R^6 and R^7 taken together complete an alicyclic or heterocyclic
system;

5   each of D^1-D^n is $(CR^6R^7)_z$ where R^6 and R^7 are as defined
above;

each of y and z is zero or an integer from 1 to 10, the
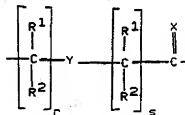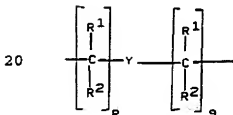sum y + z being from 2 to 10;

each of G^1-G^{n-1} is $-NR^3CO-$, $-NR^3C^S-$, $-NR^3SO-$ or $-NR^3SO_2-$,
10   in either orientation, where R^3 is as defined below;

each of A^1-A^n and B^1-B^n are selected such that:

(a) A is a group of formula (IIa), (IIb), (IIc) or
(IId), and B is N or R^3N^+; or
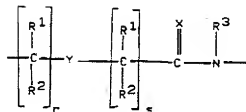
(b) A is a group of formula (IId) and B is CH;

15

20

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array} \right]_p \!\!\!-Y-\!\!\! \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array} \right]_q \qquad\qquad \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array} \right]_r \!\!\!-Y-\!\!\! \left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array} \right]_s \!\!\! \begin{array}{c} X \\ \| \\ -C- \end{array}$$

25   Formula IIa   Formula IIb

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_r Y \left[\begin{array}{c} R^1 \\ | \\ C- \\ | \\ R^2 \end{array}\right]_s \begin{array}{c} R^3 \quad O \\ | \quad \| \\ N-C- \end{array}$$

$$\left[\begin{array}{c} R^1 \\ | \\ -C- \\ | \\ R^2 \end{array}\right]_r Y \left[\begin{array}{c} R^1 \\ | \\ C- \\ | \\ R^2 \end{array}\right]_s \begin{array}{c} X \quad R^3 \\ \| \quad | \\ C-N- \end{array}$$

Formula IIc   Formula IId

wherein:

X is O, S, Se, NR^3, CH_2 or $C(CH_3)_2$;

Y is a single bond, O, S or NR^4;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, (C_1-C_4)alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen;

each R^3 and R^4 is independently selected from the group consisting of hydrogen, (C_1-C_4)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C_1-C_4)alkyl, hydroxy, alkoxy, alkylthio and amino;

Q is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NR'R''$, where R' and R'' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphos-phates, nucleotide triphos-phates, oligonucleotides, including both oligoribonucleotides

-24-

and oligodeoxyribo-nucleotides, oligonucleosides and soluble
and non-soluble polymers.

10.   A method as claimed in Claim 9, wherein said nucleic acid
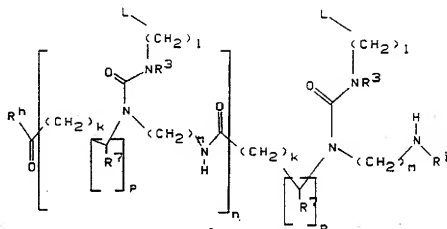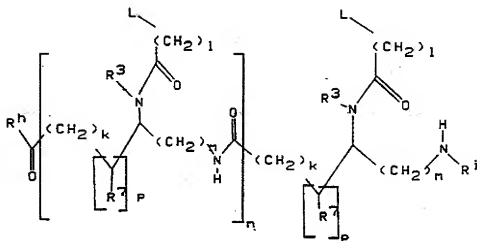analogue comprises a compound of the general formula III, IV
or V:
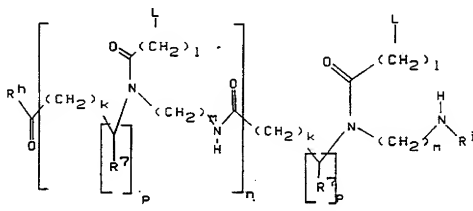
Formula III

Formula IV

Formula V

wherein:

each L is independently selected from the group consisting of hydrogen, phenyl, heterocyclic moieties, naturally occurring nucleobases, and non-naturally occurring nucleobases;

each R^7 is independently selected from the group consisting of hydrogen and the side chains of naturally occurring alpha amino acids;

-26-

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer from 1 to 5;

each p is zero or 1;

R^h is OH, NH_2 or $-NHLysNH_2$; and

R^i H or $COCH_3$.

11. A method of detecting the presence in a nucleic acid sample of a sequence, which method comprises exposing the sample of nucleic acid to a nucleic acid analogue capable of hybridising in sequence selective manner to said sequence if present in said nucleic acid sample under hybridising conditions so as to form a complex between said nucleic acid analogue and a region of said nucleic acid containing said sequence, exposing said nucleic acid to a reagent capable of degrading said nucleic acid under conditions such that the said sequence of said nucleic acid complex is protected from attack by said reagent whilst the remainder of said nucleic acid is degraded, and detecting the presence of said complex.

12. A method as claimed in claim 11, wherein the presence of said complex is detected by comparing the movement of said complex in electrophoresis with the movement of said nucleic acid analogue under similar conditions.

13. A method as claimed in claim 11 or claim 12, wherein the nucleic acid analogue bears one or more detectable labels.

14. A method as claimed in any one of claims 11 to 13, wherein said label is a fluorescent label or a radio label.

15. A method of carrying out a nucleic acid amplification comprising protecting a selected region of a nucleic acid within a nucleic acid sample by hybridising to said nucleic acid a nucleic acid analogue which hybridises thereto in a sequence specific manner, exposing said sample to a nucleic

-27-

acid attacking reagent which degrades the nucleic acid in the
sample except for said selected region, and amplifying said
selected region.

5   16. A method for working up the reaction product of a nucleic
acid amplification procedure comprising protecting a selected
region of the nucleic acid product of the amplification by
forming a complex between the nucleic acid and a nucleic acid
analogue which hybridises thereto in a sequence selective
10   manner and exposing said nucleic acid to a nucleic acid
attacking reagent, wherein said complex is more stable against
attack by said reagent than said starting nucleic acid.

17. A method as claimed in Claim 16, wherein said selected
15   region to be protected does not include at least one primer
binding site necessary for the repetition of said amplifi-
cation procedure.

18. A method as claimed in any one of Claims 11 to 17,
20   wherein the nucleic acid analogue is capable of hybridising
to a nucleic acid of complementary sequence to form a hybrid
which is more stable against denaturation by heat than a
hybrid between the conventional deoxyribo-nucleotide
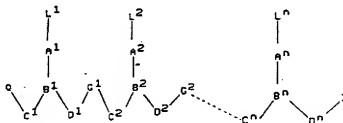corresponding in sequence to said analogue and said nucleic
25   acid.

19. A method as claimed in any one of Claims 11 to 18,
wherein the nucleic acid analogue is a peptide nucleic acid
in which said backbone is a polyamide backbone, each said
30   ligand being bonded directly or indirectly to an aza nitrogen
atom in said backbone, and said ligand bearing nitrogen atoms
mainly being separated from one another in said backbone by
from 4 to 8 intervening atoms.

35   20. A method as claimed in any one of Claims 11 to 19,
wherein the nucleic acid analogue is capable of hybridising

-28-

to a double stranded nucleic acid in which one strand has a
sequence complementary to said analogue, in such a way as to
displace the other strand from said one strand.

5   21. A method as claimed in any one of Claims 11 to 20,
wherein the nucleic acid analogue has the general formula 1:

Formula 1

wherein:

    n is at least 2,

10   each of L^1-L^n is independently selected from the group
consisting of hydrogen, hydroxy, $(C_1$-$C_4)$alkanoyl, naturally
occurring nucleobases, non-naturally occurring nucleobases,
aromatic moieties, DNA intercalators, nucleobase-binding
groups, heterocyclic moieties and reporter;

15   each of C^1-C^n is $(CR^6R^7)_y$ where R^6 is hydrogen and R^7 is
selected from the group consisting of the side chains of
naturally occurring alpha amino acids, or R^6 and R^7 are
independently selected from the group consisting of hydrogen,
$(C_2$-$C_6)$alkyl, aryl, aralkyl, heteroaryl, hydroxy, $(C_1$-$C_6)$
20   alkoxy, $(C_1$-$C_6)$alkylthio, NR^3R^4 and SR^5, where R^3 and R^4 are
as defined below, and R^5 is hydrogen, $(C_1$-$C_6)$alkyl, hydroxy,
alkoxy, or alkylthio-substituted $(C_1$ to $C_6)$alkyl or R^6 and R^7
taken together complete an alicyclic or heterocyclic system;

    each of D^1-D^n is $(CR^6R^7)_z$ where R^6 and R^7 are as defined
25   above;

each of y and z is zero or an integer from 1 to 10, the sum y + z from 2 to 10;

each of G^1-G^{n-1} is $-NR^3CO-$, $-NR^3C^5-$, $-NR^3SO-$ or $-NR^3SO_2-$, in either orientation, where R^3 is as defined below;
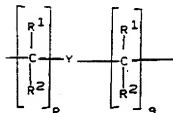
each of A^1-A^n and B^1-B^n are selected such that:

(a) A is a group of formula (IIa), (IIb), (IIc) or (IId), and B is N or R^3N^+; or

(b) A is a group of formula (IId) and B is CH;
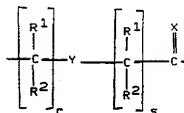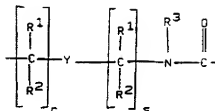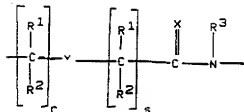
Formula IIa

Formula IIb

Formula IIc   Formula IId

wherein:

X is O, S, Se, NR^3, CH_2 or $C(CH_3)_2$;

Y is a single bond, O, S or NR^4;

each of p and q is zero or an integer from 1 to 5, the sum p+q being not more than 10;

each of r and s is zero or an integer from 1 to 5, the sum r+s being not more than 10;

each R^1 and R^2 is independently selected from the group consisting of hydrogen, (C_1-C_4)alkyl which may be hydroxy- or alkoxy- or alkylthio-substituted, hydroxy, alkoxy, alkylthio, amino and halogen; and

each R^3 and R^4 is independently selected from the group consisting of hydrogen, (c_1-C_4)alkyl, hydroxy- or alkoxy- or alkylthio-substituted (C_1-C_4)alkyl, hydroxy, alkoxy, alkylthio and amino;

Q ·is $-CO_2H$, $-CONR'R''$, $-SO_3H$ or $-SO_2NR'R''$ or an activated derivative of $-CO_2H$ or $-SO_3H$; and

I is $-NR'R''$, where R' and R'' are independently selected from the group consisting of hydrogen, alkyl, amino protecting groups, reporter ligands, intercalators, chelators, peptides, proteins, carbohydrates, lipids, steroids, nucleosides, nucleotides, nucleotide diphosphates, nucleotide triphos-

-31-

phates, oligonucleotides, including both oligoribonucleotides and oligodeoxyribonucleotides, oligonucleosides and soluble and non-soluble polymers.
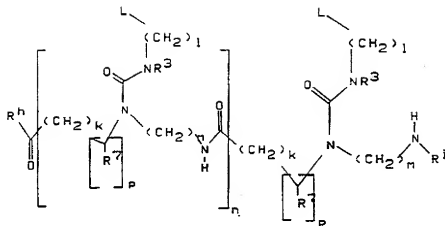
22.   A   method as claimed in Claim 21, wherein said nucleic acid analogue comprises a compound of the general formula III, IV or V:
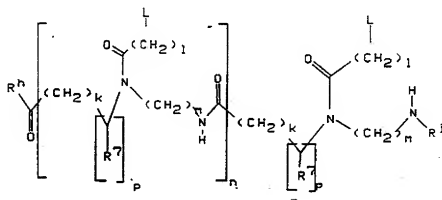
Formula III

Formula IV

wherein:

each L is independently selected from the group
consisting of hydrogen, phenyl, heterocyclic moieties,
naturally occurring nucleobases, and non-naturally occurring
nucleobases;

each R^7 is independently selected from the group
consisting of hydrogen and the side chains of naturally
occurring alpha amino acids;

n is an integer greater than 1,

each k, l, and m is, independently, zero or an integer
from 1 to 5;

each p is zero or 1;

R^h is OH, NH_2 or $-NHLysNH_2$; and

R^i is H or $COCH_3$.

23. A method of detecting a nucleic acid sequence substan-
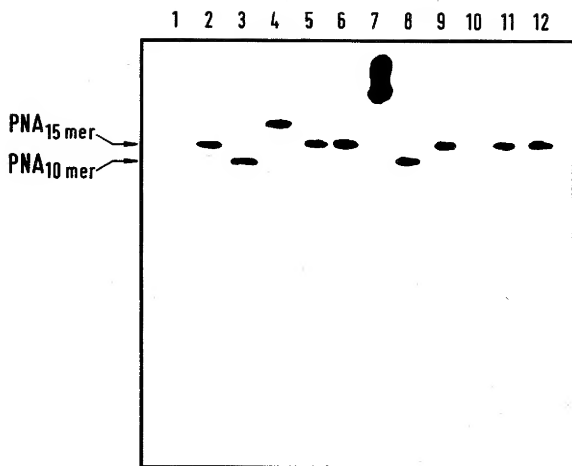tially as hereinbefore described in Example 3.

24. A method of protecting a nucleic acid from digestion by
an attacking reagent substantially as hereinbefore described
in Example 3 or Example 4 or Example 5.

-33-

25. A method of detecting a sequence variation substantially
as hereinbefore described in Example 5.

Fig. 1

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C07K14/00   C12Q1/68   C07H21/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C07H   C12Q   C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 20703 (BUCHARDT, EGHOLM, NIELSEN AND BERG.) 24 March 1995 cited in the application see the whole document -----	1-25

☐ Further documents are listed in the continuation of box C.   ☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 February 1995	2 1. 03. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Scott, J

Form PCT/ISA/210 (second sheet) (July 1992)

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9220703	26-11-92	AU-A-	1880692	30-12-92
		AU-A-	1884392	30-12-92
		BR-A-	9206049	27-12-94
		CA-A-	2109320	25-11-92
		WO-A-	9220702	26-11-92
		EP-A-	0586474	16-03-94
		EP-A-	0586618	16-03-94
		HU-A-	66597	28-12-94
		JP-T-	6509063	13-10-94
		JP-T-	6506945	04-08-94
		NO-A-	934122	11-01-94
		NO-A-	934235	20-01-94